several structural studies have demonstrated that it contains an alpha helix containing the two invariant histidine residues and two invariant cysteine residues in a beta turn co-ordinated through zinc. To date, over 10,000 zinc finger sequences have been identified in several thousand known or putative transcription factors. Zinc finger domains are involved not only in DNA-recognition, but also in RNA binding and in protein-protein binding. Current estimates are that this class of molecules will constitute about 2% of all human genes.

Please replace the paragraph beginning at line 21 of page 2 with the following rewritten paragraph:

A number of papers have reported attempts to produce ZFPs to modulate particular target sites. For example, Choo et al., *Nature* 372, 645 (1994), report an attempt to design a ZFP that would repress expression of a brc-abl oncogene. The target segment to which the ZFPs would bind was a nine base sequence 5'GCA GAA GCC3' chosen to overlap the junction created by a specific oncogenic translocation fusing the genes encoding brc and abl. The intention was that a ZFP specific to this target site would bind to the oncogene without binding to abl or brc component genes. The authors used phage display to screen a mini-library of variant ZFPs for binding to this target segment. A variant ZFP thus isolated was then reported to repress expression of a stably transfected brc-able construct in a cell line.

Please replace the paragraph beginning at line 20 of page 3 with the following rewritten paragraph:

None of the above studies provided criteria for systematically evaluating the respective merits of the different potential target sites within a candidate gene. The phage display studies by Rebar et al., supra, Jamieson et al., supra and Choo et al, *PNAS.*(1994) supra, all focused on alterations of the natural Zif268 binding-site, 5'GCG TGG GCGc3' (SEQ ID NO:11), and were not made with reference to a predetermined target gene. Choo et al. Nature (1994), supra's selection of target site was constrained solely by the intent that the site overlap the interface between brc and abl segments and did not involve a comparison of different potential target sites. Likewise, Greisman & Pabo chose certain target sites because of their known regulatory roles and did not consider the relative merits of different potential target segments within a preselected target gene. Similarly, Choo et al. (1998), supra's choice of target site within ras was constrained by the position of a mutation. No criterion is provided for Choo et al. (1998)'s selection of a target site in HIV Tat. Finally, both Pomerantz et al., supra and Liu

a3 et al., supra constructed artificial hybrid target sites for composite zinc fingers and then inserted the target sites into reporter constructs.

Please replace the paragraph beginning at line 26 of page 9 with the following rewritten paragraph:

a4 Fig. 2 shows a three finger zinc finger protein bound to a target site (SEQ ID NO:12) containing three D-able subsites.

Please replace the paragraph beginning at line 28 of page 14 with the following rewritten paragraph:

a5 Linkage can be accomplished using any of the following peptide linkers. TGEKP (SEQ ID NO:2) (Liu et al., 1997, supra.); $(G_4S)_n$ (SEQ ID NO:3) (Kim et al., *PNAS* 93, 1156-1160 (1996.); GGRRGGGS (SEQ ID NO:4); LRQRDGERP (SEQ ID NO:5); LRQKDGGGSERP (SEQ ID NO:6); LRQKD$(G_3S)_2$ERP (SEQ ID NO:7). Alternatively, flexible linkers can be rationally designed using computer program capable of modeling both DNA-binding sites and the peptides themselves or by phage display methods . In a further variation, noncovalent linkage can be achieved by fusing two zinc finger proteins with domains promoting heterodimer formation of the two zinc finger proteins. For example, one zinc finger protein can be fused with fos and the other with jun (see Barbas et al., WO 95/119431).

Please replace the paragraph beginning at line 12 of page 15 with the following rewritten paragraph:

a6 A component finger of zinc finger protein typically contains about 30 amino acids and has the following motif (N-C) (SEQ ID NO:1):

Cys-$(X)_{2-4}$-Cys-X.X.X.X.X.X.X.X.X.X.X.X-**His**-$(X)_{3-5}$-His

-1 1 2 3 4 5 6 7

Please replace the paragraph beginning at line 24 of page 15 with the following rewritten paragraph:

a7 The process of designing or selecting a nonnaturally occurring or variant ZFP typically starts with a natural ZFP as a source of framework residues. The process of design or selection serves to define nonconserved positions (i.e., positions -1 to +6) so as to confer a desired binding

specificity. One suitable ZFP is the DNA binding domain of the mouse transcription factor Zif268. The DNA binding domain of this protein has the amino acid sequence:

(F1) YACPVESCDRRFSRSDELTRHIRIHTGQKP

(F2) FQCRICMRNFSRSDHLTTHIRTHTGEKP

(F3) FACDICGRKFARSDERKRHTKIHLRQK (SEQ ID NO:8)

and binds to a target 5' GCG TGG GCG 3'.

Please replace the paragraph beginning at line 1 of page 16 with the following rewritten paragraph:

Another suitable natural zinc finger protein as a source of framework residues is Sp-1. The Sp-1 sequence used for construction of zinc finger proteins corresponds to amino acids 531 to 624 in the Sp-1 transcription factor. This sequence is 94 amino acids in length. The amino acid sequence of Sp-1 is as follows

PGKKKQHICHIQGCGKVYGKTSHLRAHLRWHTGERP

FMCTWSYCGKRFTRSDELQRHKRTHTGEKK

FACPECPKRFMRSDHLSKHIKTHQNKKG (SEQ ID NO:9)

Sp-1 binds to a target site 5'GGG GCG GGG3'.

Please replace the paragraph beginning at line 9 of page 16 with the following rewritten paragraph:

An alternate form of Sp-1, an Sp-1 consensus sequence, has the following amino acid sequence:

meklrngsgd

PGKKKQHACPECGKSFSKSSHLRAHQRTHTGERP

YKCPECGKSFSRSDELQRHQRTHTGEKP

YKCPECGKSFSRSDHLSKHQRTHQNKKG (SEQ ID NO:10) (lower case letters are a leader sequence from Shi & Berg, *Chemistry and Biology* 1, 83-89. (1995). The optimal binding sequence for the Sp-1 consensus sequence is 5'GGGGCGGGG3'. Other suitable ZFPs are described below.

Please replace the paragraph beginning at line 7 of page 23 with the following rewritten paragraph:

In the formula 5'NNx aNy bNzc3', the triplets of NNx aNy and bNz represent the triplets of bases on the target strand bound by the three fingers in a zinc finger protein. The complements of the highlighted bases are the sites of potential fourth base binding on the nontarget strand. If only one of x, y and z is a G, and this G is followed by a K, the target site includes a single D-able subsite. For example, if only x is G and a is K, the site reads **NNG KNy** bNz w with the D-able subsite highlighted. If both x and y but not z are G and a and b are K, then the target site has two overlapping D-able subsites as follows: 5'**NNG *KNG* *K*Nz** c3' (SEQ ID NO:13) with one such site being represented in bold and the other in italics. If all three of x, y and z are G and a, b and c are K, then the target segment includes three D-able subsites, as follows 5'**NNG *KNG* <u>KNG K</u>**3' (SEQ ID NO:14), the D-able subsites being represented by bold, italics and underline.

Please replace the paragraph beginning at line 2 of page 44 with the following rewritten paragraph:

GNGGNNGNN(N){0,3}**GNGG**NNGNNN (SEQ ID NOS:15 and 16)

Please replace the paragraph beginning at line 3 of page 44 with the following rewritten paragraph:

GNGGNNGNN(N){0,3}GNN**GNGG**NNN (SEQ ID NOS:17 and 18)

Please replace the paragraph beginning at line 4 of page 44 with the following rewritten paragraph:

GNGGNNGNN(N){0,3}GNGGNN**GNGG** (SEQ ID NOS:19 and 20)

Please replace the paragraph beginning at line 5 of page 44 with the following rewritten paragraph:

GNN**GNGG**NN(N){0,3}**GNGG**NNGNNN (SEQ ID NOS:21 and 22)

Please replace the paragraph beginning at line 6 of page 44 with the following rewritten paragraph:

GNN**GNGG**NN(N){0,3}GNN**GNGG**NNN (SEQ ID NOS:23 and 24)

Please replace the paragraph beginning at line 7 of page 44 with the following rewritten paragraph:

a16 GNN**GNGG**NN(N){0,3}GNGGNN**GNGG** (SEQ ID NOS:25 and 26)

Please replace the paragraph beginning at line 8 of page 44 with the following rewritten paragraph:

a17 GNNGNN**GNGG**(N){0,3}**GNGG**NNGNNN (SEQ ID NOS:27 and 28)

Please replace the paragraph beginning at line 9 of page 44 with the following rewritten paragraph:

a18 GNNGNN**GNGG**(N){0,3}GNN**GNGG**NNN (SEQ ID NOS:29 and 30)

Please replace the paragraph beginning at line 10 of page 44 with the following rewritten paragraph:

a19 GNNGNN**GNGG**(N){0,3}GNGGNN**GNGG** (SEQ ID NOS:31 and 32)

Please replace the paragraph beginning at line 11 of page 44 with the following rewritten paragraph:

a20 GNNGNN**GNGGNGG**NNGNNN (SEQ ID NO:33)

Please replace the paragraph beginning at line 12 of page 44 with the following rewritten paragraph:

a21 GNNGNN**GNGG**NN**GNGG**NNN (SEQ ID NO:34)

Please replace the paragraph beginning at line 13 of page 44 with the following rewritten paragraph:

a22 GNNGNN**GNGG**NNGNN**GNGG** (SEQ ID NO:35)

Please replace the paragraph beginning at line 25 of page 44 with the following rewritten paragraph:

a23 KNGGNNKNN(N){0,3}KNGGNNKNNN (SEQ ID NOS:36 and 37)

Please replace the paragraph beginning at line 26 of page 44 with the following rewritten paragraph:

a24 KNGGNNKNN(N){0,3}KNNKNGGNNN (SEQ ID NOS:38 and 39)

Please replace the paragraph beginning at line 27 of page 44 with the following rewritten paragraph:

a25 KNGGNNKNN(N){0,3}KNNKNNKNGG (SEQ ID NOS:40 and 41)

Please replace the paragraph beginning at line 28 of page 44 with the following rewritten paragraph:

a26 KNNKNGGNN(N){0,3}KNGGNNKNNN (SEQ ID NOS:42 and 43)

Please replace the paragraph beginning at line 1 of page 45 with the following rewritten paragraph:

a27 KNNKNGGNN(N){0,3}KNNKNGGNNN (SEQ ID NOS:44 and 45)

Please replace the paragraph beginning at line 2 of page 45 with the following rewritten paragraph:

a28 KNNKNGGNN(N){0,3} KNNKNNKNGG (SEQ ID NOS:46 and 47)

Please replace the paragraph beginning at line 3 of page 45 with the following rewritten paragraph:

a29 KNNKNNKNGG(N){0,2}KNGGNNKNNN (SEQ ID NOS:48 and 49)

Please replace the paragraph beginning at line 4 of page 45 with the following rewritten paragraph:

a30 KNNKNNKNGG(N){0,2}KNNKNGGNNN (SEQ ID NOS:50 and 51)

Please replace the paragraph beginning at line 5 of page 45 with the following rewritten paragraph:

a31 KNNKNNKNGG(N){0,2}KNNKNNKNGG (SEQ ID NOS:52 and 53)

Please replace the paragraph beginning at line 6 of page 45 with the following rewritten paragraph:

a32 KNNKNNKNGGNGGNNKNNN (SEQ ID NO:54)

Please replace the paragraph beginning at line 7 of page 45 with the following rewritten paragraph:

a33 KNNKNNKNGGNNKNGGNNN (SEQ ID NO:55)

Please replace the paragraph beginning at line 8 of page 45 with the following rewritten paragraph:

a34 KNNKNNKNGGNNKNNKNGG (SEQ ID NO:56)

Please replace the paragraph beginning at line 14 of page 45 with the following rewritten paragraph:

a35 KNGKNNKNN(N){0,3}KNGKNNKNNN (SEQ ID NOS:57 and 58)

Please replace the paragraph beginning at line 15 of page 45 with the following rewritten paragraph:

a36 KNGKNNKNN(N){0,3}KNNKNGKNNN (SEQ ID NOS:59 and 60)

Please replace the paragraph beginning at line 16 of page 45 with the following rewritten paragraph:

a37 KNGKNNKNN(N){0,3}KNNKNNKNGK (SEQ ID NOS:61 and 62)

Please replace the paragraph beginning at line 17 of page 45 with the following rewritten paragraph:

Please replace the paragraph beginning at line 18 of page 45 with the following rewritten paragraph:

a38 KNNKNGKNN(N){0,3}KNNKNGKNNN (SEQ ID NOS:65 and 66)

Please replace the paragraph beginning at line 19 of page 45 with the following rewritten paragraph:

a39 KNNKNGKNN(N){0,3}KNNKNNKNGK (SEQ ID NOS:67 and 68)

Please replace the paragraph beginning at line 20 of page 45 with the following rewritten paragraph:

a40 KNNKNNKNGK(N){0,2}KNGKNNKNNN (SEQ ID NOS:69 and 70)

Please replace the paragraph beginning at line 21 of page 45 with the following rewritten paragraph:

a41 KNNKNNKNGK(N){0,2}KNNKNGKNNN (SEQ ID NOS:71 and 72)

Please replace the paragraph beginning at line 22 of page 45 with the following rewritten paragraph:

a42 KNNKNNKNGK(N){0,2}KNNKNNKNGK (SEQ ID NOS:73 and 74)

Please replace the paragraph beginning at line 23 of page 45 with the following rewritten paragraph:

a43 KNNKNNKNGKNGKNNKNNN (SEQ ID NO:75)

Please replace the paragraph beginning at line 24 of page 45 with the following rewritten paragraph:

a44 KNNKNNKNGKNNKNGKNNN (SEQ ID NO:76)

Please replace the paragraph beginning at line 25 of page 45 with the following rewritten paragraph:

a45 KNNKNNKNGKNNKNNKNGK (SEQ ID NO:77)

Please replace the table beginning at line 4 of page 47 ("Table 3") with the following rewritten table:

a46 Table 3

TARGET NAME	SEQUENCE	PROTEIN NAME	Kd (nM)	SEQ ID NO:
FAD 1	**GAG GTA GAG G**	FAD 1A	10	78
FAD 1	**GAG GTA GAG G**	FAD 1B	10	78
FAD 2	GTC **GTG TGG** A	FAD 2A	100	79
FAD 3	GTT **GAG GAA** G	FAD 3A	100	80
FAD 3	GTT **GAG GAA** G	FAD 3B	100	80
FAD 4	**GAG GTG GAA** G	FAD 4A	10	81
FAD 4	**GAG GTG GAA** G	FAD 4B	2	81
FAD 5	**TAG GTG GTG** A	FAD 5A	10	82

Please replace the paragraph beginning at line 30 of page 48 with the following rewritten paragraph:

a47

The 22 ZFPs designed to targets with two GG type D-able subsites have the strongest binding affinity with an average Kd = 15 nM. Of the 50 ZFPs having a Kd < 100 nM, 49 have at least one D-able subsite. The table shows the following conclusion: (1) binding to a target site with one D-able subsite bind more strongly than ZFPs binding to a target site lacking a D-able subsites; (2) ZFPS binding to a target site with two D-able subsites bind more strongly than ZFPs that bind to a target sing with one D-able subsite, and (3) ZFPs with a target site with a GG D-able subsite bind more strongly than ZFPs with a target site with a GT D-able subsite.

Please replace the paragraph beginning at line 27 of page 53 with the following rewritten paragraph:

a48

(If the subsite is of the form xxA, xxC or xxT, its score also remains unchanged.)

Please replace the paragraph beginning at line 8 of page 54 with the following rewritten paragraph:

a49

(When using this option, the program considers the identity of base immediately to the 3' side of the target site (in lower case). For the last site, this base is undefined in this example and this is noted by placing the pound sign '#' at this position.)

Please replace the paragraph beginning at line 22 of page 55 with the following rewritten paragraph:

a50

Triplet	3	2	1	F1	F2	F3	Finger SEQ ID NO:
[1]	5'TGCGGGGCA			********	********	*ERDHLRT	88
[3]	5'GGGGCGGGG			********	*RSDELQR	********	89
[4]	5'GAGTGTGTG			*RKDSLVR	********	********	90

DISORDERED:

********	*RSDELTR[2](3)	********	91
********	*RSDERKR[2](1)	********	92

Please replace the paragraph beginning at line 32 of page 55 with the following rewritten paragraph:

The 'ordered' output shows that, in the ZFP data table, there is one instance where the TGC subsite is contacted by a zinc finger in the third triplet of a target site. The finger in this case is ERDHLRT (SEQ ID NO:88), and the site is 5'TGCGGGGCA3'. There is also one similar instance for each of the other two subsites - GCG, and GTG. The fingers in these cases are, respectively, RSDELQR (SEQ ID NO:89) and RKDSLVR (SEQ ID NO:90). This information is used to propose the three finger protein F1-RKDSLVR, F2-RSDELQR, F3-ERDHLRT (SEQ ID NO:93) as a design to bind the target 5'TGCGCGGTG3'.

Please replace the paragraph beginning at line 6 of page 56 with the following rewritten paragraph:

The 'disordered' output shows that there are two cases in the ZFPdata table in which fingers contact a GCG subsite, but not at the center subsite of a target. Rather, in one case GCG is contacted at the 5' end, and the other the 3' end, and in these cases the finger sequences are RSDELTR (SEQ ID NO:91) and RSDERKR (SEQ ID NO:92). These are alternate designs for binding GCG in the target site.

Please replace the table beginning at line 15 of page 58 ("Table 9") with the following rewritten table:

Table 9: Exemplary ZFP data table

#	target site	ZFP sequence F1	F2	F3	reference information	ZFP SEQ ID NO:
1	TGCGGGGCA	RSADLTR	RSDHLTR	ERDHLRT	SBS design GR-223, Kd: 8 nM	94
2	GCGTGGGCG	RSDELTR	RSDHLTT	RSDERKR	Zif 268, Kd: 0.04 nM	95
3	GGGGCGGGG	KTSHLRA	RSDELQR	RSDHLSK	SP1, Kd: 25 nM	96
4	GAGTGTGTG	RKDSLVR	TSDHLAS	RSDNLTR	SBS design GL-8.3.1, Kd: 32 nM	97

Please insert the accompanying paper copy of the Sequence Listing, page numbers 1 to 33, at the end of the application.

In the claims: